The Durham Children’s Data Center
January 3, 2015

Executive Summary

At a time when public resources are scarce, policy decisions regarding investments in children and families must take into account rigorous evidence about the likely effectiveness of those investments. With public accountability as a goal, the Durham community establishes the Durham Children’s Data Center. Initial partners in the Data Center include the Durham County Manager’s Office, the Durham Public Schools, the Durham Partnership for Children, and Duke University. The Data Center will be governed by a team of Durham community leaders and will be co-led initially by Durham Public Schools Superintendent Bert L’Homme, Assistant Durham County Manager Drew Cummings, and Duke University Professors Kenneth Dodge and Clara Muschkin.

The Data Center will collaborate with Durham community leaders to identify policy questions for which the community seeks answers, identify administrative data files that can be examined, and then complete research reports for community leaders. In addition to ongoing informal relationships, the Data Center will bring together community leaders and university faculty members in periodic public assemblies, at which policy questions will be generated, reports of research findings will be presented and discussed, and new research studies will be planned.

Under the direction of the Data Center leadership, and to the extent allowed by law and policy, a Data Center Director will receive and combine individual-level administrative data files from Durham County (including the public schools, early childhood programs, healthcare, labor, welfare programs, and income and property taxes) with data files from other sources and levels (such as the U.S. Census and criminal justice) toward the goal of data analysis to inform policy decisions. Duke University faculty members will support rigorous analysis of data for the dual goals of answering questions that are of immediate and direct interest to Durham community leaders and providing research reports for the broader academic and policy communities about children’s needs. With matched administrative data files, the Data Center will provide a descriptive analysis of how government resources are distributed across families, how these resources relate to children’s status over time, and how limited community resources might be deployed to improve children’s outcomes such as kindergarten readiness and health.

Initial funding for the Data Center will be provided by Duke University. The Center will be housed and administered at the Sanford School of Public Policy, and faculty members from all schools within Duke will participate. Founding community participants will include the Durham Public Schools and the Durham County Manager’s Office, and additional public agencies and community organizations will be welcome to participate.
Background

Over the past several years, child and education policy has moved toward center stage in domestic political discourse. Important policy discussions are emerging at the federal, state, and local levels. Educators, community leaders, and elected officials are searching for scientific evidence and hard data to guide public policy. Decision-makers want to know which programs work, whether programs are cost-beneficial, and what new topics should receive their attention. In response to the challenge of making evidence-based public policy decisions, Durham community leaders have initiated the Durham Children’s Data Center to bring scientific empirical analysis to bear on policy decision-making for the Durham, North Carolina, community.

Public policy for children and families has reached its zenith of public attention. In education, questions are being asked about standardized testing, school choice, impact of teachers, and kindergarten readiness. In health, questions are being asked about Medicaid expansion, benefit-cost ratios for preventive care, and impact of prenatal care. In welfare, questions are being asked about work-family balance, childcare quality, protective services impact, and assimilation of immigrant families. Even more questions need to be asked about common factors that might underlie issues in all of these domains, such as income disparities and parenting effectiveness. However, most of what we know about these topics has been studied in siloed domains with little cross-fertilization of ideas. The Durham Children’s Data Center will bring together data files across diverse sectors for the purpose of understanding how government resources affect trajectories of child development in disparate areas of health, education, and welfare.

Durham community leaders have expressed interest in rigorous analysis of data to inform their decision-making, and they have turned to Duke University for support. Skills of scientific inquiry, empirical modeling of data, and accurate interpretation of findings are ones that university scholars bring to the policy table. However, these skills are often used for academic scholarship rather than direct application in policy decision-making. The Data Center is an innovative collaboration between community leaders and university faculty members to support policy making for Durham’s children and families.

Durham public leaders would like their decisions to be more informed by empirical analysis of the administrative data that they collect and control. These data files include individual-level information about a child’s early childhood program participation, government subsidies, healthcare, childcare, foster care, and education. However, these administrative data files have not been merged to address important policy questions. Because public agencies do not have sufficient resources to conduct the analyses that they want, they have initiated a collaboration with the Sanford School of Duke University.

The Data Center will benefit children and families in Durham by providing evidence to inform policy making. The Data Center will also bring benefits to the university community through experiences of understanding how policy is made, learning the priorities of local government, and interacting with policy makers. Duke faculty members will contribute to the Durham community by providing expertise in conducting analyses of data on Durham’s children and in writing reports for Durham community leaders. In addition to these benefits and contributions, faculty members at the Sanford School, upon review and assurance of appropriate data security safeguards, will receive access to the confidential data files for academic research that goes beyond immediate policy questions to address topics of interest to the broader academic and policy communities. Duke students will have opportunities to learn by participating in all phases of the research process, from formulation of policy research questions to the planning of a research design, and from analysis of data to application of findings to new policy.

The Durham Children’s Data Center aspires to become a national model for how a community and university can collaborate to place knowledge in service to society.
Project Description

The Durham Children’s Data Center will have three core components.

**Posing policy questions.** A critical first step in policy analysis is the formulation of an answerable research question. This step involves posing a broad question, framing it in quantifiable terms, refining the question so that it can answered with data, and prioritizing its components. Some policy issues can be informed by descriptive analysis of existing data. Other issues require evaluation of programs and policies by following children across time. Still other questions, particularly regarding causation, cannot be answered fully unless rigorous research designs are implemented.

The first component of the Data Center, therefore, will be a dynamic “Think Tank” to generate important policy questions which leaders in the Durham community want answered. The major work of the Data Center will be driven by the questions that community leaders pose. They will identify questions and prioritize work, and Sanford faculty members will help them frame their questions in answerable ways. Durham leaders are interested in questions such as how to maximize government resources to improve children’s kindergarten readiness, how to raise the population literacy rate among Durham’s third graders, and how to reduce disparities in children’s outcomes across groups defined by gender, race, and income. Urgent challenges include addressing the needs of immigrant families and families stressed by job loss.

This component of the Data Center will be implemented through a series of meetings between Durham community leaders and Sanford faculty members, convened by the Data Center Director. Some of these meetings will occur informally, and at least one meeting annually will be a public event at which Durham’s major policy issues for children and families will be posed and discussed.

The most pressing and important policy research questions will be formulated and prioritized during this stage, and a plan will be written for each question that will include the operationalization of the research question, budget and personnel needed, one or more Sanford faculty member leaders, and a timeline for completion.

**Acquiring and analyzing data to answer policy questions.** A core task of the Data Center will be the assembly of administrative data files to answer the important policy questions posed by Durham leaders. Durham government and nonprofit participants will provide access to data files, including confidential individual-level data on children, in order to answer the questions that they pose. These data files are collected and stored primarily to fulfill government requirements for documentation of program implementation, billing, or consumer service. Although these data files exist in government agencies, they have rarely been used to evaluate program impact or inform future policy. Even rarer is their use for generation of knowledge.

As important questions are posed by community leaders, the need for access to particular data files becomes clear. Some questions require individually-identified data files that must be linked across sectors (such as linking health records and education records to answer whether children with particular healthcare needs are at academic risk, or linking childcare records with education records to answer whether childcare subsidies improve education outcomes). Other questions might not require identities of children but information about their context, such as a neighborhood or school. Data files exist at many levels, from individuals tracked across time to institutions such as schools and police districts. The Data Center Director will receive files and inspect data codebooks and characteristics to determine what modifications and resources will be needed to use data to address questions. As success with initial data files is demonstrated, other county files may be accessed, such as delinquency and crime data. It might also be possible to merge county data with other files, such as children’s food environment information.
Some policy research questions require sophisticated evaluation designs (such as a randomized controlled trial or longitudinal design) or new data collection (such as a survey, qualitative interviews, or ethnographic observation). Sanford faculty members and community leaders will interact to determine what data and evaluation designs are possible within resource and policy constraints. Once data files are accessed, a data analyst will clean and merge files, write codebooks, work with Sanford faculty members to analyze data, and draft research reports for Durham community leaders.

A crucial issue for the Data Center is data security. Health records, education records, and welfare records are extremely sensitive, and their access is governed by carefully drawn policies and procedures (HIPPA, FERPA, etc.). Fortunately, the North Carolina Education Research Data Center, housed at the Duke Center for Child and Family Policy, has 15 years of experience in storing sensitive education and health data files, setting policy for access, following legal and professional procedures for management of confidential data, following university Institutional Review Board (IRB) policies, and interacting with government agencies for data exchange. The Sanford School’s Director of Informational Technology is quite experienced with university policies about data security to ensure that Sanford’s data security plans conform with university policy. The Data Center will draw on this experience in setting and implementing policies for data security.

**Communication of findings for policy.** The third stage of policy analysis to communicate research findings in a coherent manner that answers the policy question driving the study. Sanford faculty members and the Data Center Director will work with Sanford communications experts to draft a written product for each question that will ordinarily include a long technical report, a several-page lay summary, and a brief executive summary.

Some reports will be confidentially delivered to Durham community leaders. Some reports will be communicated through a media release following Durham community leaders’ guidance. Some reports will be communicated through PowerPoint and oral presentations to targeted audiences. The findings of each study will be communicated to the Durham leader(s) who issued the request. Reports for scholarly audiences will be communicated in advance to Durham community leaders and will be disseminated through academic journals and policy briefs.

**Additional benefit to Duke’s academic research mission**

Although the central function of the Data Center will be to serve the Durham community, Durham leaders will allow the data files to be used for academic research by Duke faculty members and students, as long as policies are followed for data privacy and use. Each research study will be proposed to the Data Center and reviewed by a team that will include representatives of the Durham agencies from which data will be drawn and Duke University. Each proposed research study will be reviewed and approved in advance by the Duke IRB, the Duke Office of Information Technology for data-security review, and relevant Durham agency leaders. Agency leaders can veto any study in advance but cannot censor findings once a study has been approved.

**Additional benefit to student learning**

Whenever possible, Duke undergraduate, masters, and doctoral students will participate in the activities of the Data Center, through independent studies; formal coursework; honors, masters, and doctoral theses; and research assistantships. All student involvement will be proposed to the Data Center governing body for prior approval.
Evaluation of Success and Impact
The success and impact of the Data Center will be realized in how local policy making is altered. National impact and visibility will be enhanced through communication about the Data Center itself and the research reports that are generated.

Durham Community Leaders

Michael Becketts, Director, Durham Division of Social Services
Laura Benson, Executive Director, Durham Partnership for Children
Heidi Carter, Chair, Durham Public Schools Board
Suzanne Cotterman, Director of Early Education, Durham Public Schools
Drew Cummings, Assistant Durham County Manager
Wendell Davis, Durham County Manager
Teresa Daye, Executive Director, Curriculum, Instruction, & Assessment, Durham Public Schools
Gayle Harris, Director, Durham County Department of Health
Bert L’Homme, Superintendent, Durham Public Schools
Ann Oshel, Chief Community Relations Officer, Alliance Behavioral Health
Max Rose, Program Manager, MDC
Julie Spencer, Assistant Superintendent for Research and Accountability, Durham Public Schools
Stacey Wilson-Norman, Deputy Superintendent for Academics, Durham Public Schools
Phail Wynn, Vice President, Durham and Regional Affairs, Duke University

Duke Faculty Members

Elizabeth Ananat          Sanford
Paul Bendich             Mathematics
Lawrence Carin           Engineering
Charles Clotfelter        Sanford
Philip Cook              Sanford
Sandy Darity             Sanford
Kenneth Dodge            Sanford
Helen Egger              Psychiatry
Gavan Fitzsimons          Fuqua
Anna Gassman-Pines       Sanford
Christina Gibson-Davis   Sanford
Amar Hamoudi             Sanford
Matt Harding             Sanford
Nicole Heilbron          Psychiatry
Sunshine Hillygus        Political Science
Rick Hoyle               Psychology and Neuroscience
Sunny Ladd               Sanford
Clara Muschkin           Sanford
Candice Odgers           Sanford
Marcos Rangel            Sanford
Seth Sanders             Sanford
Guillermo Sapiro         Engineering
Staff Members

Yu Bai, Senior Data Analyst
Elizabeth Gifford, Director
Barbara Pollock, Financial Officer
Kendra Rose, Clinical Advisor
Brandi Thomas, Staff Assistant

Oversight

The Durham Children’s Data Center will be governed by a team of Durham community leaders and Duke University faculty members. An executive committee will include Durham community leaders Bert L'Homme and Drew Cummings and Duke University Professors Kenneth Dodge and Clara Muschkin.

The Center will be administered at the Sanford School of Public Policy.